Peer-reviewed papers

Original Road Safety Research

First-Stage Evaluation of a Prototype Driver Distraction Human-Machine-Interface Warning System

Christine Mulvihill[1], Tim Horberry[1], Michael Fitzharris[1], Brendan Lawrence[1], Raphaela Schnittker[1], Mike Lenné[2], Jonny Kuo[2], Darren Wood[3]

[1] *Monash University Accident Research Centre, Clayton, Victoria, Australia*
[2] *Seeing Machines Ltd, Fyshwick, ACT, Australia*
[3] *Ron Finemore Transport Ltd, Wodonga, Victoria, Australia*

Corresponding Author: Christine Mulvihill, Building 70, 21 Alliance Lane, Monash University, Victoria, 3800. christine.mulvihill@monash.edu Tel: 03 9905 4367

This peer-reviewed paper was first submitted as an Extended Abstract and an Oral Presentation was recommended by two reviewers at the 2021 Australasian Road Safety Conference (ARSC2021) held online, 28-30 September 2021. The two Reviewers also recommended that the Extended Abstract be expanded into a 'Full Paper' and undergo further peer-review as a journal submission by three independent experts in the field. The Extended Abstract is published in the ARSC2021 Proceedings with a link guiding readers to this 'Full Paper' version which is being reproduced here with the kind permission of the authors and will only be available in this edition of the JRS.

Key Findings

- Four variations of the Human Machine Interface (HMI) of a multi-modal driver distraction warning system developed by the authors were evaluated in a truck simulator;
- Driver acceptance of the HMIs was assessed using the System Acceptance Scale; and salience, comprehension and perceived effectiveness of components of the HMIs (modality, intensity of warning) were assessed using likert scales;
- Participants accepted the HMIs and understood the warning components and considered these to be effective;
- Further testing is recommended to validate the findings and assess their safety impact on road.

Abstract

Recent advances in vehicle technology permit the real-time monitoring of driver state to reduce distraction-related crashes, particularly within the heavy vehicle industry. Relatively little published research has evaluated the human machine interface (HMI) design for these systems. However, the efficacy of in-vehicle technology depends in large part on the acceptability among drivers of the system's interface. Four variations of the HMI of a prototype multi-modal warning system developed by the authors for driver distraction were evaluated in a truck simulator with eight car drivers and six truck drivers. Driver acceptance of the HMIs was assessed using the System Acceptability Scale; and salience, comprehension and perceived effectiveness of components of the HMIs (modality, intensity of warning) were assessed using likert scales. The results showed that participants considered the HMIs to be acceptable and useful, and that the warning components were largely noticed, understood correctly, and perceived to be effective. Although this study identified no major design flaws with the recently developed HMIs, further simulator testing with a larger sample size is recommended to validate the findings. On-road evaluations to assess the impact of the HMIs on real world safety are a necessary pre-requisite for implementation.

Received: 01/08/2021; **Received in revised form:** 13/09/2021; **Accepted:** 13/09/2021;

Available on-Line: 10/11/2021 https://doi.org/10.33492/JRS-D-21-00049

Keywords

Driver Distraction, Human Machine Interface, Warning Systems, Simulation, In-Vehicle Technology.

Introduction

Inattention is a significant contributing factor to crashes and occurs when a driver fails to pay sufficient attention to activities that are required for safe driving (Lee, Young & Regan, 2009; Regan, Hallet & Gordon, 2011). Data from the Australian National Crash In-depth Study found that driver inattention contributed to 63% of serious casualty crashes in which a known crash contributing factor was identified (Beanland, Fitzharris, Young & Lenné, 2013). In 25% of these crashes the driver was distracted by a competing activity, most commonly a non-driving related activity within the vehicle. Distraction has been identified as a major challenge in the commercial vehicle safety context particularly given the high mileage that truck drivers routinely cover and the greater severity of injuries associated with heavy vehicle collisions (BITRE, 2019).

Alongside regulation and enforcement, recent advances in vehicle technology permit the real-time monitoring of driver state to mitigate driver impairment (e.g., Lenné et al., 2018). Evidence suggests that driver monitoring systems (DMS) are a promising approach for detecting driver distraction (e.g., Liang, Reyes, & Lee, 2007; Miyaji, Kawanaka & Oguri, 2009) as well as other forms of driver inattention including fatigue (e.g., Fitzharris, Lui, Stephens & Lenné, 2017).

From an end-user perspective, relatively little published research has evaluated the Human Machine Interface (HMI) design for the warning component of DMS. This represents an important knowledge gap since the efficacy of any in-vehicle technology depends as much on its acceptance by the driver as it does on optimal system performance (van der Laan, Heino & de Waard, 1997; Regan, Stevens & Horberry, 2014). Research also suggests that drivers with favourable opinions of DMS tend to experience more or greater safety benefits than drivers with unfavourable attitudes (e.g., Blanco et al., 2009).

Although acceptance has been variously defined in the scientific literature (see Regan et al., 2014 for a full review), the term can be broadly defined as the degree to which a driver perceives the benefits derived from a system as greater than the costs (Blanco et al., 2009). This is important because technology is likely to be ignored or even circumvented if its safety benefit is outweighed by factors such as annoyance, mistrust or lack of effectiveness.

A handful of HMI design evaluations for warning systems conducted under driver distraction have examined various dimensions of user acceptance including: perceived usefulness, trust, annoyance and distraction (Anund & Nilsson, 2020; Biondi, Strayer, Rossi, Gastaldi & Mulatti, 2017; Donmez, Boyle & Lee, 2007; Kircher, Kircher & Ahlstrom, 2009; Kujala, Karvonen & Mäkelä, 2016; Lee, Hoffman & Hayes, 2004; Maltz & Shinar, 2007; Roberts, Ghazizadeh & Lee, 2012). Collectively, the results of these studies suggest generally high levels of driver acceptance of the warning systems. However, while perceived usefulness and trust were rated highly in most of the studies, drivers still rated the systems as annoying or distracting in about half of the studies.

Three of these evaluations included comparisons of different warning types. Lee et al. (2004) found that cognitively distracted drivers were more trusting of multi-stage than single stage visual-tactile alerts. Maltz and Shinar (2007) found that drivers perceived multi-stage auditory alarms to be more annoying and less helpful than single stage auditory alarms when drivers were distracted by a visual-manual task. Biondi et al. (2017) found that multimodal auditory-tactile warnings were no more frustrating than single modality auditory or tactile warnings when drivers were cognitively distracted. Less is known about drivers' perceptions of specific sub-components of warnings such as modality type (visual, haptic, auditory), warning intensity, and timing of the warnings. These factors contribute to the overall acceptance of warnings and are important in optimising the design of the system for final implementation. In the optimisation process, understanding the extent to which drivers notice and understand the warnings is critical. This is particularly important in the commercial vehicle context where drivers must attend to many other technologies present in modern day truck cabs. Surprisingly, there has been little research done in this area for driver distraction.

A tertiary goal of the Advanced Safe Truck Concept project (Lenné et al., 2017) was to develop a prototype driver distraction warning system. The prototype system was developed by the authors using human centred design (HCD) principles (Horberry et al., 2021) and issued warnings to drivers when they were distracted by a visual-manual secondary task. A multi-stage iterative process was adopted to design the prototype system. This included a comprehensive review of literature and HMI design guidelines; interviews and design workshops with truck drivers and stakeholders; HMI evaluation studies and finalising the HMI concepts (See Horberry et al., 2021, for full details of this process). As part of this, an evaluation method was formalised and is described in this paper.

Using the adopted evaluation methods, this paper provides an initial examination of four variations of the HMI of the prototype driver distraction warning system. The HMI was a two-level system that comprised a cautionary warning (level 1) followed by an urgent warning (level 2) if distraction persisted after the initial alert. The system was multimodal with simultaneously presented auditory alerts at both levels comprising 'beep' tones, a spoken warning and tactile vibration through the seat. A visual alert was also issued at the urgent (i.e., level 2) warning stage.

The primary aim was to determine whether drivers considered each of the four HMIs to be acceptable and to determine which one was most effective for alerting the driver. A secondary aim was to determine whether drivers understood and considered effective the various components of the HMIs (modality, intensity of warning) so that refinements could be made for future testing. Although the warning system was developed primarily for use in a commercial vehicle safety context, both car and truck drivers were tested given the likely applicability of the system across a range of vehicle types.

Methods

Participants

Fourteen drivers were included in the study. Eight were car drivers recruited from the Monash University Accident Research Centre (MUARC) Participant Database and six were truck drivers recruited from the general population. The mean age of the drivers was 38.4 years and three were female (21.5%). The drivers had an average of 19 years' driving experience (15 years for car drivers, and nearly 24 years for truck drivers). None reported having a diagnosed hearing impairment while one truck driver reported being colour blind. Drivers reported that they generally listened to the radio when driving and had a wide range of experience with in-vehicle technology.

Car drivers were paid $50 for their participation. Truck drivers were paid $100 as they were required to answer additional questions following the simulator drive. The study was approved by the Monash University Human Research Ethics Committee (Project Number: 8247).

Interface conditions

Two escalation timings for the driver distraction HMI prototype were tested between presentation of the first (cautionary) and second (urgent) level warnings: short-escalating (1.5 seconds from cautionary to urgent warning) and long-escalating (4 seconds from cautionary to urgent warning). The cautionary (level 1) and urgent (level 2) warnings were also tested separately in addition to a 'no warnings' (control) condition. The order of presentation of the interface conditions was counterbalanced between participants, with the exception of the control which was always presented third to prevent drivers expecting a warning after every distraction event. The four interface conditions are described in Table 1. Full details of the warning specifications can be found in Horberry et al. (2021).

System Acceptance Scale (SAS) and perceived effectiveness of interface conditions

Participants' subjective experience of the four interface conditions was measured using the System Acceptance Scale (SAS) (van der Laan, Heino & de Waard, 1997). The SAS comprised nine questions on two dimensions: usefulness and affective satisfying. Participants provided responses on 5-point likert scales from -2 to 2, with zero being the mid-point and higher scores indicating positive ratings. Overall system acceptance was the mean score on the usefulness and satisfying dimensions.

Table 1. Interface conditions tested

Interface condition	Description/device components
Control	No warning issued
Cautionary (Level 1)	Auditory alerting tone (70 dBA) 500 m/s later, spoken message, Australian female voice (70 dBA) 'Pay attention' Tactile warning (5 pulses in driver's seat, duty cycle 1), triggered at the same time as the speech message
Urgent (Level 2)	Auditory alerting tone (90 dBA) 350 m/s later, spoken message, Australian female voice (70 dBA) 'Pay attention!' Tactile warning (5 pulses in driver's seat, duty cycle 0.3), triggered at the same time as the speech message Visual warning: eye on road graphic as a Head Up Display (HUD) on the windscreen, triggered at the same time as the tactile and speech message
Short-escalating	The cautionary warning (level 1) followed 1.5 seconds later by the urgent warning (level 2)
Long-escalating	The cautionary warning (level 1) followed 4 seconds later by the urgent warning (level 2)

Participants were asked to rate how effective was the multimodal warning presentation for each interface condition on a four-point likert scale with: 4 very effective, 3 effective, 2 ineffective and 1 very ineffective. They were also asked to rank the four interface conditions in order of their effectiveness for alerting the driver (1 being most effective and 4 being least effective).

Comprehension

To assess whether participants understood the various modalities within the four interface conditions, the experimenter asked, ‘What does the audio/speech/vibration/visual warning mean?’ and ‘What do the two different levels of warning mean?’ Comprehension was assessed by the experimenter and measured on a scale from 0-2, where 0 represented no or incorrect understanding, 1 represented partial understanding and 2 represented full understanding.

Perceived effectiveness

To measure participants’ ratings of perceived effectiveness of the various modalities within the four interface conditions the experimenter asked i) ‘How effective was the audio/speech/vibration/visual warning in altering you?’ overall and then separately for each level across the four interface conditions (where applicable) and ii) ‘How effective was presenting all of the different warning modalities at the same time?’ Participants were also asked to rate how effective was the time gap between the first and second level warnings in the short-escalating and long-escalating conditions. These were all measured on a four-point scale with: 4 very effective, 3 effective, 2 ineffective and 1 very ineffective.

Salience

To assess whether participants noticed the four interface conditions, the experimenter asked ‘What warnings did you notice?’ This question always preceded administration of the SAS and the comprehension and perceived effectiveness questions to ensure that participants were aware of the warnings they were being asked to evaluate and could provide valid assessments.

Post-drive open questions

Participants were asked a general question about what they thought of each of the interface conditions and to provide any further views about the HMI overall.

MUARC advanced truck driving simulator

The experiment was conducted using the Advanced Truck Driving Simulator at MUARC, which comprised a full-size Volvo truck cab with automatic transmission, a real steering wheel, brake and accelerator pedals, engine noise and low frequency vibration to simulate a running engine and cabin vibration. The simulated driving scenario was projected onto a 180-degree cylinder forward screen and

Figure 1. Truck simulator exterior with forward road scene projection

a flat rear screen. Figure 1 shows the outside of the truck simulator with the front image projected.

Distraction task

Participants completed a distraction task four times throughout the drive. The distraction task was a self-paced text messaging task implemented with a 7-inch Samsung Galaxy Tab A touch screen tablet computer (Model Number SM-T280) mounted on the centre console. The distraction task required participants to read pre-generated text messages that contained phases with one word missing and type in the missing word(s) (or “pass”) to complete the phrase. This task was selected to induce visual and manual input from the driver which have been shown to be significant sources of distraction when driving.

Procedure

Participants were tested individually in MUARC’s advanced truck driving simulator. Upon arrival, the study was explained to participants after which they provided informed consent and completed a questionnaire designed to capture demographic and driving experience data.

Before driving, and whilst seated in the truck looking at the forward roadway, participants received each of the four interface conditions presented in counterbalanced order. Participants provided feedback after each interface condition using the salience questions and the SAS.

Participants then completed a 20-minute drive in the truck simulator. The drive took place on a single carriageway country road at between 80-100 km/h with infrequent traffic. Approximately five minutes into the drive, participants were instructed to engage in the distraction task (about 30 seconds) and were then presented with one of the four interface conditions. Participants were informed that they should ‘respond in the way that you think is appropriate’ after each warning. About 15 seconds after the warning had ceased, they were asked to stop the truck and provide feedback using the salience, comprehension and perceived effectiveness questions. This cycle was repeated four times (including the control, i.e., no warning) with each interface condition presented in the same counterbalanced order as per the pre-drive. At the conclusion of the drive, each interface condition was again played in turn whilst the participant was still seated in the truck, and feedback was given using the SAS. Participants then completed the post-drive open questions and were paid for their participation.

Data analysis

To test for differences across the different interface conditions and modalities, participant responses were analysed using IBM SPSS Statistics (Version 26). Given the small sample size, non-parametric testing was used. Friedman's test was used for comparisons of acceptance and effectiveness. Where significant differences were found, Wilcoxon signed-ranked tests for multiple comparisons were used with a Bonferroni correction applied to the alpha level. Wilcoxon signed-ranked tests were used to compare ratings of effectiveness between the cautionary and urgent levels within each of the audio tone, audio speech and tactile modalities. Statistical significance was set at $p \leq 0.05$.

Results

Comparisons between the four interface conditions

Salience

Prior to administration of the SAS participants were asked 'What warnings did you notice?' The warnings were issued a second time if participants reported that they did not notice one or more components. During the drive, between 1-2 participants reported not noticing the auditory tone in each condition while one participant reported not noticing the speech warning and one reported not noticing the visual warning.

Acceptability

The mean scores obtained for usefulness, satisfying and acceptance for each interface condition on the SAS are shown in Table 2. All systems were considered useful, as evidenced by all the scores being strongly positive on the usefulness dimension. In contrast, the scores for the satisfying dimension were mostly neutral or slightly negative, with only the cautionary condition achieving a positive rating. Overall acceptance was rated positively across all conditions, although slightly less so than for usefulness.

The mean ratings for the satisfying dimension differed significantly across the four conditions $\chi2(3) = 17.182$, $p = 0.001$. Post-hoc tests revealed that the cautionary condition was rated as significantly more satisfying than each of the other conditions: $Z_{long\text{-}escalating} = -3.049$, $p = 0.002$; $Z_{short\text{-}escalating} = -2.982$, $p = 0.003$; $Z_{urgent} = -2.831$, $p = 0.005$. The four interface conditions did not differ significantly on usefulness $\chi2(3) = 4.514$, $p > 0.05$, or overall acceptance $\chi2(3) = 8.951$, p = $p > 0.05$.

Perceived effectiveness

For each interface condition, participants were asked to rate how effective was the multimodal warning presentation (i.e., presenting all the warnings at the same time).

Across the four conditions the majority of participants rated the multimodal warning presentation as effective or very effective (see Figure 2). Overall, 21% of ratings were ineffective or very ineffective because participants thought the warnings created too much sensory overload or were too startling. The mean effectiveness ratings (See Table 3) did not differ significantly across the four conditions $\chi2(3) = 2.824$, $p > 0.05$.

Participants were asked to rank the four interface conditions in order of effectiveness from 1 being the most effective to 4 being the least effective. The ranking sum, and the mean, median and mode for the rankings are shown in Table 4, along with the number and percentage of participants who rated each condition as most effective.

Table 4 shows that, overall, the condition regarded as most effective was the short-escalating condition, followed by the long-escalating condition (with the cautionary and urgent conditions being equally ranked lower). However, the rankings were not significantly different $\chi2(3) = 3.794$, $p > 0.05$.

Table 2. Scores on usefulness, satisfying and acceptance for the four interface conditions

Metric Mean (SD)	**Cautionary**	**Urgent**	**Short-escalating**	**Long-escalating**
Usefulness	0.97 (0.93)	1.46 (0.61)	1.43 (0.59)	1.26 (0.74)
Satisfying	0.70 (1.0)	-0.21 (1.2)	-0.07 (1.1)	1.26 (0.74)
Acceptance	0.83 (0.78)	0.62 (0.79)	0.68 (0.68)	0.37 (0.77)

Scores ranged from -2 to 2.

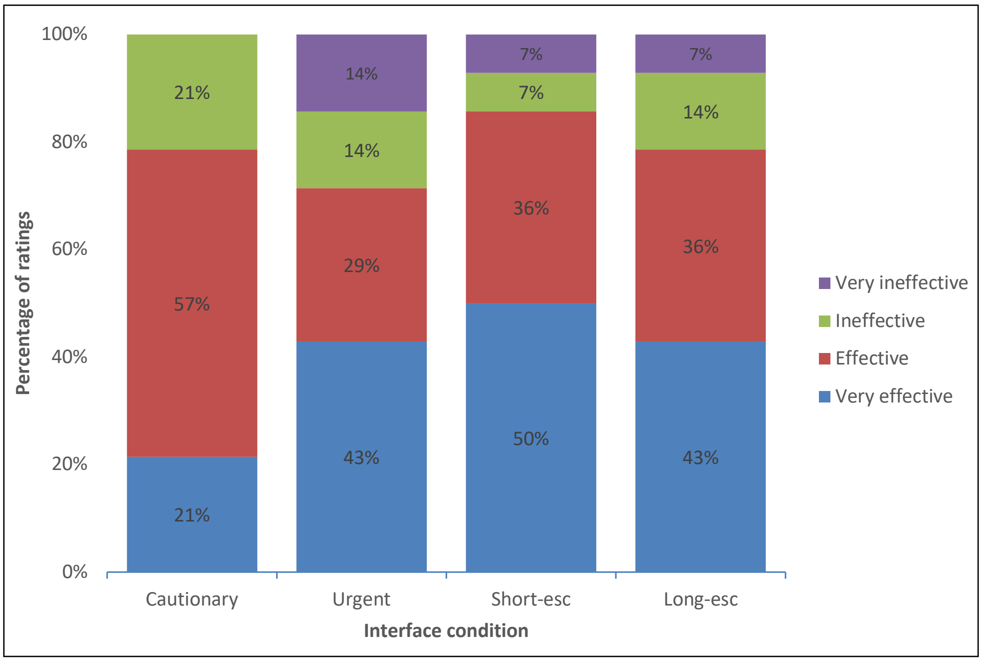

Figure 2: Percentage of effectiveness ratings for the four interface conditions

Comparison between warning modalities (audio, tactile and visual)

Salience

Participants were asked what warnings they noticed. The warnings were issued a second time if participants failed to notice one or more components. Across each of the four conditions, between 2-3 participants did not report hearing the cautionary tone warning or the urgent tone warning, and between 1-2 participants missed the speech, vibration or visual warnings.

Participants were asked whether and to what extent any background noise affected their ability to notice the warnings. All participants except one said that background noise did not affect the noticeability of the warnings.

Comprehension

Table 3. Mean effectiveness ratings (multimodal presentation) for the four interface conditions

	Cautionary	Urgent	Short-escalating	Long-escalating
Mean	3	3	3.28	3.14

Table 4. Rankings of perceived effectiveness for the four interface conditions

	Cautionary	Urgent	Short-escalating	Long-escalating
Ranking sum*	40	40	30	33
Mean ranking	2.86	2.86	2.14	2.36
Median ranking	4	3	2	2
Mode	4	3	1	1
Number of drivers who ranked as most effective	4	1	5	4
Percent of drivers who ranked as most effective	28.6%	7.1%	35.7%	28.6%

*Lower number means a higher ranking

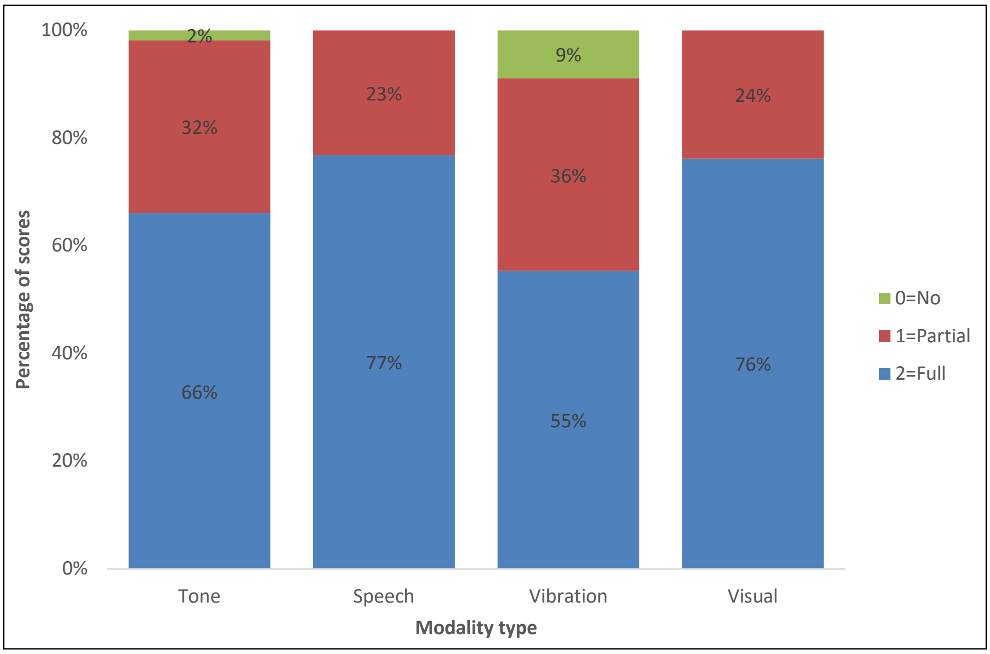

Figure 3. Percentage of no, partial and full understanding scores for the different warning modalities

For each of the four conditions participants were asked to state their understanding of each of the warning modalities - tone, speech, seat vibration, and visual warning. Figure 3 presents the percentage of comprehension scores rated as no (0), partial (1) or full (2) understanding for each modality averaged across the four conditions.

Most participants achieved full understanding across each of the four warning modalities (Figure 2). This ranged from 55% (vibration) to 77% (speech). All participants understood – either partially or fully - the audio speech and the visual warnings, while one did not understand the audio tone (i.e., scored zero) (2%) and five did not understand the seat vibration (9%). Those who did not understand the seat vibration warning thought it was a seat belt warning or audio tactile line markings.

Participants were asked to state their understanding of what the two different levels of warning meant in the short-escalating and long-escalating conditions. In the long-escalating condition, 71.4% of participants had a full understanding of the intent of the warning while the remaining 28.6% indicated partial understanding. For the short-escalating warning, 42.9% had a full understanding of the intent of the warning and 57.1% indicated partial understanding.

Perceived effectiveness

Participants were asked to provide a rating of effectiveness using a 4-point scale ranging from 1 (very ineffective) to 4 (very effective) for each of the two levels of warning separately, and then another rating for the warning overall 'as a package'. Figure 4 shows the percentage of effectiveness ratings by modality type averaged across the four warning conditions.

Participants were more likely to rate the speech warnings (Level 2 and Overall) and the Level 2 vibration warning as 'very effective' compared to the other modalities. All remaining modalities were more likely to be rated as 'effective'. The visual warning was more likely to be rated as 'ineffective' (31%) or 'very ineffective' (14%) than the other modalities. All modalities were considered slightly more effective in the escalated state, (noting that the visual warning was not used for a cautionary warning), however none of the differences were significant ($p > 0.05$).

There was a significant difference in perceived effectiveness across the four warnings ($\chi2(3) = 9.637$, $p = 0.022$) (See Table 5). Post-hoc tests showed that none of the differences were significant at the stringent Bonferroni alpha level of 0.0008, although the overall speech and the overall tone warnings were rated as significantly more effective than the visual warning at the 0.05 alpha level: $Z_{Speech} = -2.550$, $p = 0.011$; $Z_{Tone} = -2.227$, $p = 0.026$.

Participants were asked to rate how effective was the timing from the first to the second level warning. The majority rated the time gap in both conditions as effective, although this was slightly higher in the short-escalating condition (57.1%) than the long-escalating condition (50%).

Post drive open questions

All 14 participants provided feedback about the four interface conditions. In all responses the main comment was that the warnings were effective in attracting the driver's attention. Eleven participants (78%) indicated that the cautionary warning was not as noticeable as the other warnings and five (35%) noted the urgent warning was startling or a 'bit sudden' compared to the other warnings. Three participants (21%) thought the urgent or long-

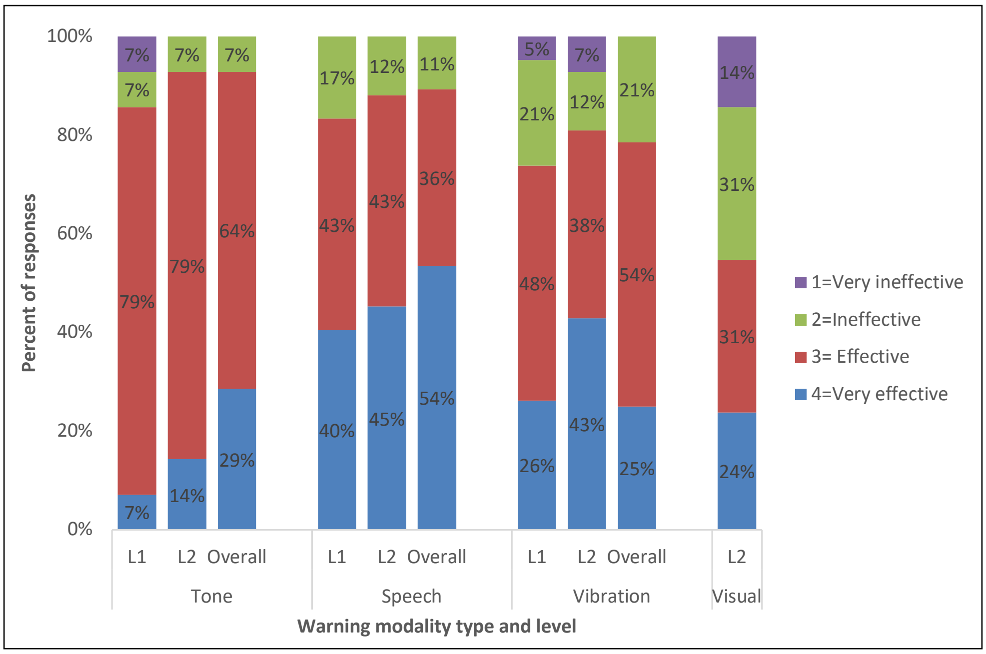

Figure 4: Percentage of effectiveness ratings for the different warning levels and modalities

escalating warnings were annoying, particularly in relation to the cautionary warning.

Five participants (35%) made mention of the two-stage conditions being especially useful for attracting attention because they provided more time for the driver to respond and were not as startling as the urgent (single level) warning. For example, '*It is better to have the warnings build up over time and with the shorter (1.5s) gap in between*' and '*the urgent warning is least preferred as it happens too suddenly*'. Two of these participants thought the gap should be reduced in the long-escalating condition while one thought it should be increased in the short-escalating condition. Conversely, two drivers thought the two-level warnings were more annoying than the single level warnings: '*My preference is for single level warnings because they are not as annoying but the multiple level warnings were more alerting and so they are probably more effective*'.

Table 5: Mean effectiveness ratings and SDs for the different warning levels and modalities

	Level	Mean & SD
Modality		
Tone	Level 1	2.93 (0.40)
	Level 2	3.07 (0.32)
	Overall	3.21 (0.51)
Speech	Level 1	3.18 (0.62)
	Level 2	3.31 (0.56)
	Overall	3.43 (0.61)
Vibration	Level 1	2.95 (0.64)
	Level 2	3.17 (0.85)
	Overall	3.03 (0.66)
Visual	Level 2	2.59 (0.91)

One participant disliked the multimodal warning presentations for the reason that '*There are already a lot of warnings especially on the dashboard and it could end up startling you so that you do something unsafe like run off the road*'. Conversely, another participant thought the warning redundancy was useful because '*if you miss one warning you'll notice another one*'.

Three participants (21%) gave positive feedback about the tactile warnings which were regarded as preferable for their unique alerting potential. *'The seat vibration is good because it is unlike any other warning'*. Two participants (14%) thought the visual warning could be distracting, particularly if drivers were drawn to look at it rather than at the road, while three participants thought the visual warning could easily be overlooked if the driver was looking away from the road. Three drivers thought the auditory speech warning was too loud/starling and two thought the auditory tone was too weak. Overall, no major issues or consistent design deficiencies were reported.

The truck drivers were asked if they thought the warnings could interfere with other in-cab devices. Only the truck drivers were asked this question due to their having direct exposure to DMS technology as part of their regular employment. Of the five drivers who answered this question, two thought the warnings might interfere with other in-cab devices because *'there are already a lot of warnings in the cabin and lots happening to distract you (especially the auditory warnings)'.* Two drivers who disagreed stated that *'all the warnings are different to each other'* and consequently there was little perceived risk of interference with other in-cab devices. One driver gave a more neutral view in which interference was thought to occur only in specific situations: *'Potentially if, for example, a low oil warning alert goes off at the same time as the distraction warning'.* No major interference from the HMI with other in-cab systems was therefore noted, overall.

Discussion

The primary aim of this study was to determine whether drivers considered each of the four HMI conditions to be acceptable and to determine which one was most effective for alerting the driver. A secondary aim was to determine whether drivers understood and considered effective the various components of the HMI conditions (modality, intensity of warning, etc.) so that refinements could be made for future testing. The four HMI conditions tested in this study were accepted by the participant group that included both experienced truck and car drivers. Although the short-escalating system was ranked as the most effective of the four HMIs, the differences were not significant. Assessment of the different warning modalities demonstrated that each one was generally noticed, correctly understood and considered effective. Equally, they were largely perceived as not likely to excessively interfere with other in-cab warnings in a truck.

As the HMI was a warning system, it is unsurprising that the scores for the satisfying dimension were neutral or slightly negative, despite an overall high level of system acceptance of the four HMIs. The cautionary warning was the only one to receive a positive score on this dimension and was rated as being significantly more satisfying than the other three warning conditions. This is likely because the cautionary warning was the least intrusive of the four warnings; it was single level, had fewer sensory modalities (i.e. no visual warning) and was issued at a lower audible volume and tactile vibrational frequency. Notably, however, some level of annoyance may be necessary in achieving the alerting function of a warning system. For example, some participants suggested that the cautionary warning was the only one that was not regarded as being overly alerting. Indeed, previous studies have also found that, despite an overall high level of acceptance of distraction warning systems, drivers also found them annoying and/or distracting (Kujala et al., 2016; Lee et al., 2004; Maltz & Shinar, 2007; Roberts et al., 2012).

Although the short-escalating system scored highest for usefulness and was ranked as the most effective of the four HMIs, the differences were not significant. A larger sample size may be necessary to verify the current findings.

All systems were generally well detected during the drive despite the simulated engine noise and cabin vibration designed to represent a typical truck use environment. Although up to three participants failed to notice one or more warning modalities during the drive, only one participant stated that background noise affected their ability to detect the vibration warning. Nonetheless, these findings underscore the importance of warning redundancy in that if one warning modality is missed, other modalities still should be detected. Previous studies have shown that multimodal warning presentations were more effective than unimodal warning presentations in reducing visual and cognitive distraction because redundancy provides an additional means of attracting the driver's attention (Biondi et al., 2017; Ho, Reed & Spence, 2007). This is supported by the results of the current study in which the majority of participants rated 'presenting all the warnings at the same time' as being effective or very effective for each condition. Four participants gave a rating of ineffective or very ineffective for at least one of the four interface conditions because they thought the warnings were too startling or could have potential to overload a driver. Thus, a balance needs to be achieved between providing adequate warning for the driver but not overloading or startling them from the use of multimodal presentations.

The auditory and speech modalities in the current warnings were perceived to be more effective than the visual warning, although these differences failed to reach statistical significance once the stringent Bonferroni adjustment for multiple comparisons was applied. A larger sample size would likely provide a better test of any real differences between warning modalities. Previous research has shown that visual warnings were not effective for attracting the attention of drivers who were looking away from the road (Campbell, Richard, Brown & McCallum, 2007) and were also ineffective, or not as effective, as collision warnings presented in the auditory or tactile modalities for attentive drivers (e.g., Campbell et al., 2007; Scott & Gray, 2008; Kiefer et al., 1999). These findings are also consistent with a number of comments made by participants that the visual warning could easily be overlooked if the driver was looking away from the road, or could be distracting, particularly if drivers were drawn to look at the warning itself rather than at the road. Future research should test the effectiveness of each of the warnings without the visual component. This would also provide a means of testing whether there is a reduction in the potential for drivers to feel overloaded and/or startled by presenting multiple warnings at once, as identified by a small proportion of participants in this study.

About a third of all drivers stated that they preferred the two-level warnings because they provided more time for the driver to respond. The two-level warnings were also thought to be beneficial because their intensity built up over time, unlike the single level urgent warning which was perceived by some to be too sudden. The gaps between the two-level warning systems were rated as being equally effective. However, some participants suggested that the four second gap was too long while the 1.5 second gap was too short to allow sufficient time for drivers to correct their behaviour and then re-orient their attention back to the road. Notably, the shorter gap was also less likely to be fully understood compared to the longer gap. Future research could test the efficacy of a two-second gap between the first and second level warning. This period is consistent with the average time it takes drivers to 'safely' avert their eyes from the road without deviating from their lane and/or experiencing a collision.

Conclusions

Overall, this study found that participants accepted the HMIs and understood the warning components and considered these to be effective. Further simulator testing with a larger sample size is recommended to validate the findings, particularly with respect to driver acceptability of the visual warning and the relative effectiveness of the two multi-level warning interfaces. On-road testing to assess the impact of the HMI on real world safety is also a necessary pre-requisite for implementation.

Acknowledgements

We acknowledge the funding received from the Commonwealth of Australia through the Cooperative Research Centre Projects scheme and the cash and/or in kind support provided by Seeing Machines Ltd, Monash University, Ron Finemore Transport and the Volvo Group Australia. We also thank the participants who took part in this study.

References

Anund, A. & Nilsson, L. (2020). *ADAS&ME Project Final Report. Deliverable 10.4.* Retrieved from: https://www.adasandme.com/wp-content/uploads/2020/04/D10.4-Project-Final-Report.pdf

Ayoob, E.M., Steinfeld, A., & Grace, R. (2003). Identification of an appropriate drowsy driver interface for commercial vehicle operations. *Proceedings of the Human Factors and Ergonomics Society, 47th Annual Meeting, 47*(16),1840-1844. doi.org/10.1177/154193120304701602

Beanland, V., Fitzharris, M., Young, K.L., & Lenné, M.G. (2013). Driver inattention and driver distraction in serious casualty crashes. Data from the Australian National Crash In-depth Study. *Accident Analysis & Prevention, 54,* 99-107.

Biondi, F., Strayer, D.L., Rossi, R., Gastaldi, M., & Mulatti, C. (2017). Advanced driver assistance systems: Using multimodal redundant warnings to enhance road safety. *Applied Ergonomics, 58,* 238-244. DOI: 10.1016/j.apergo.2016.06.016

Blanco, M., Docanegra, J., Morgan, J., Fitch, G., Medina, A., Olson, R.L., Hanowski, R., Daily, B.,…Green, K. (2009). *Assessment of a drowsy-driver warning system for heavy-vehicle drivers. Final Report.* (DOT HS 811 117). Washington: US Department of Transportation, NHTSA. Retrieved from https://www.nhtsa.gov/sites/nhtsa.gov/files/811117.pdf

Boer, E. R., Rakauskas, M. E., Ward, N. J., & Goodrich, M. A. (2005). Steering entropy revisited. *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 25-32. doi.org/10.17077/drivingassessment.1139

Briggs, G. F., Hole, G. J., & Land, M. F. (2011). Emotionally involving telephone conversations lead to driver error and visual tunneling. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(4), 313-323. doi: 10.1016/j.trf.2011.02.004.

Bureau of Infrastructure, Transport and Regional Economics (BITRE). (2019). *Road trauma involving heavy vehicles 2017 crash statistical summary.* Retrieved from https://www.bitre.gov.au/sites/default/files/HV%20Annual%202017%20Mar2019_II.pdf

Campbell, J. L., Richard, C. M., Brown, J. L., & McCallum, M. (2007). *Crash warning system interfaces: Human factors insights and lessons learned.* Report (No. DOT HS 810 697). Washington, DC, USA: National Highway Traffic Safety Administration. Retrieved from https://www.nhtsa.gov/sites/nhtsa.gov/files/cws_hf_insights_task_5_final_rpt.pdf

Donmez, B., Boyle, L., & Lee, J. (2007). Safety implications of providing real time feedback to distracted drivers. *Accident Analysis & Prevention, 39,* 581□590.

Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*(2), 97-120. doi: 10.1016/j.trf.2005.04.012.

Fitzharris, M., Liu, S., Stephens, A., & Lenné, M.G. (2017). The relative importance of real-time in cab and external feedback in managing fatigue in real world commercial transport operations. *Traffic Injury Prevention, 81*, S71-78.

Ho, C., N. Reed, N. & Spence, C. (2007). Multisensory in-car warning signals for collision avoidance. *Journal of the Human Factors and Ergonomics Society, 49*(6), 107–1114.

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brown, J. (2006). Driver distraction: the effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis & Prevention, 38*(1), 185-191. doi: 10.1016/j.aap.2005.09.007.

Horberry, T., Mulvihill, C., Fitzharris, M., Lawrence, B., Lenné, M., Kuo, J., & Wood, D. (2021). Human-Centered Design for an in-vehicle truck driver fatigue and distraction warning system. *IEE Transactions on Intelligent Transport Systems.* doi: 10.1109/TITS.2021.3053096.

Kiefer, R., LeBlanc, D., Palmer, M., Salinger, J., Deering, R., & Shulman, M. (1999). *Development and validation of functional definitions and evaluation procedures for collision warning/avoidance systems.* Report (No. DOT HT 808 964). Washington, DC, USA: National Highway Traffic Safety Administration.

Kircher, K., Kircher, A., & Ahlstrom, C. (2009). *Results of a field study on a driver distraction warning system.* Report (VTI 639A). Linköping Sweden, VTI. Retrieved from https://www.diva-portal.org/smash/get/diva2:675374/FULLTEXT02.pdf

Kujala T., Karvonen, H., & Mäkelä J. (2016). Context sensitive distraction warnings–Effects on drivers' visual behavior and acceptance. *International Journal of Human Computer Studies, 90*, 39□52.

Lee, J.D., Hoffman, J.D., & Hayes, E. (2004). Collision warning design to mitigate driver distraction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vienna, Austria, 6*(1), 65-72.

Lee, J.D., Young, K.L., & Regan, M.A. (2009). Defining driver distraction. In M.A. Regan, J.D. Lee, K.L. Young (Eds), *Driver distraction: Theory, effects and mitigation* (1st ed., pp. 31 - 40). CRC Press.

Lenné, M. G., Fitzharris, M., Horberry, T., Wood, D., da Cruz, J., Kuo, J., … Peden, M. (2017). *The Advanced Safe Truck Concept (ASTC) Project: Defining and developing future concepts to enhance heavy vehicle safety.* 1-2. Abstract from Australasian Road Safety Conference 2017, Perth, Australia. https://acrs.org.au/files/papers/arsc/2017/Lenne_00176_EA.pdf

Lenné, M., Kuo, J., Fitzharris, M., Horberry, T., Mulvihill, C., Blay, K., … Peden, M., (2018). *The Advanced Safe Truck Concept Project: A partnership program developing future integrated driver monitoring technology.* In A. Hržić (Ed.), 25th ITS World Congress, Copenhagen, Denmark, 17-21 September 2018 (pp. 1-8). ERTICO - ITS Europe, Brussels, Belgium.

Liang, Y., Reyes, M. L., & Lee, J. D. (2007). Real-time detection of driver cognitive distraction using support vector machines. *IEEE Transactions on Intelligent Transportation Systems, 8*(2), 340-350.

Maltz, M., & Shinar, D. (2007). Imperfect in-vehicle collision avoidance warning systems can aid distracted drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 10* (4), 345-357. doi: https://doi.org/10.1016/j.trf.2007.01.002.

Miyaji, M., Kawanaka, H., & Oguri, K. (2009). Driver's cognitive distraction detection using physiological features by the AdaBoost. *12th International IEEE Conference on Intelligent Transportation Systems*, 1-6. doi: 10.1109/ITSC.2009.5309881.

Regan, M.A., Hallett, C., & Gordon, C.P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis and Prevention 43*, 1771–1781.

Regan, M. A., Stevens, A., & Horberry, T. (2014). Driver acceptance of new technology: Overview. In M. A. Regan, T. Horberry, & A. Stevens (Eds.), *Driver acceptance of new technology: Theory, measurement and optimisation* (pp. 3-8). Ashgate.

Roberts, S. C., Ghazizadeh, M., & Lee, J. D. (2012). Warn me now or inform me later: Drivers' acceptance of real-time and post-drive distraction mitigation systems. *International Journal of Human-Computer Studies, 70*(12), 967-979. doi: https://doi.org/10.1016/j.ijhcs.2012.08.002.

Scott, J.J., & Gray, R. (2008). A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. *Journal of the Human Factors and Ergonomics Society, 50(*2), 264–275.

van der Laan, J.D., Heino, A., & de Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research - Part C: Emerging Technologies,* 5, 1-10.